


Nama: Ibnu Fajar Setiawan NIM: 065002000006	 Praktikum Data Warehouse	MODUL 2 Nama Dosen: Ir. Teddy Siswanto, MMSi
Hari/Tanggal: Hari, 28-03-2022		Nama Asisten Labratorium: 1. Azhar Rizki Zulma 065001900001 2. Nadiya Amanda Rizkania 064001900003

Input Data pada Spoon

1. Teori Singkat

Data warehouse adalah jenis sistem manajemen data yang dirancang untuk memungkinkan dan mendukung kegiatan business intelligence (BI), terutama analitik. Gudang data semata-mata dimaksudkan untuk melakukan kueri dan analisis dan sering berisi sejumlah besar data historis. Data dalam gudang data biasanya berasal dari berbagai sumber seperti file log aplikasi dan aplikasi transaksi. Gudang data memusatkan dan mengkonsolidasikan sejumlah besar data dari berbagai sumber. Kemampuan analitisnya memungkinkan organisasi untuk memperoleh wawasan bisnis yang berharga dari data mereka untuk meningkatkan pengambilan keputusan. Seiring waktu, ia membangun catatan sejarah yang dapat sangat berharga bagi para ilmuwan data dan analisis bisnis. Karena kemampuan ini, gudang data dapat dianggap sebagai "sumber kebenaran tunggal" organisasi.

2. Alat dan Bahan

Hardware : Laptop/PC

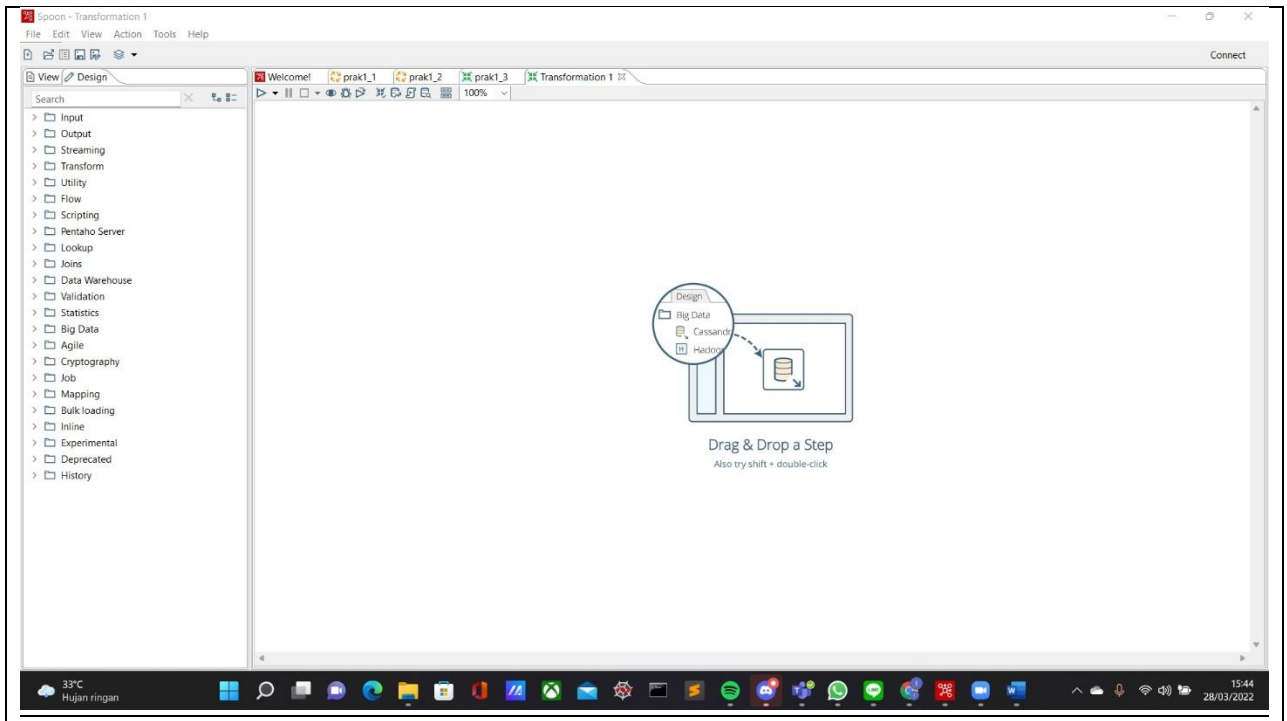
Software : Spoon Pentaho from Hitachi Vantara



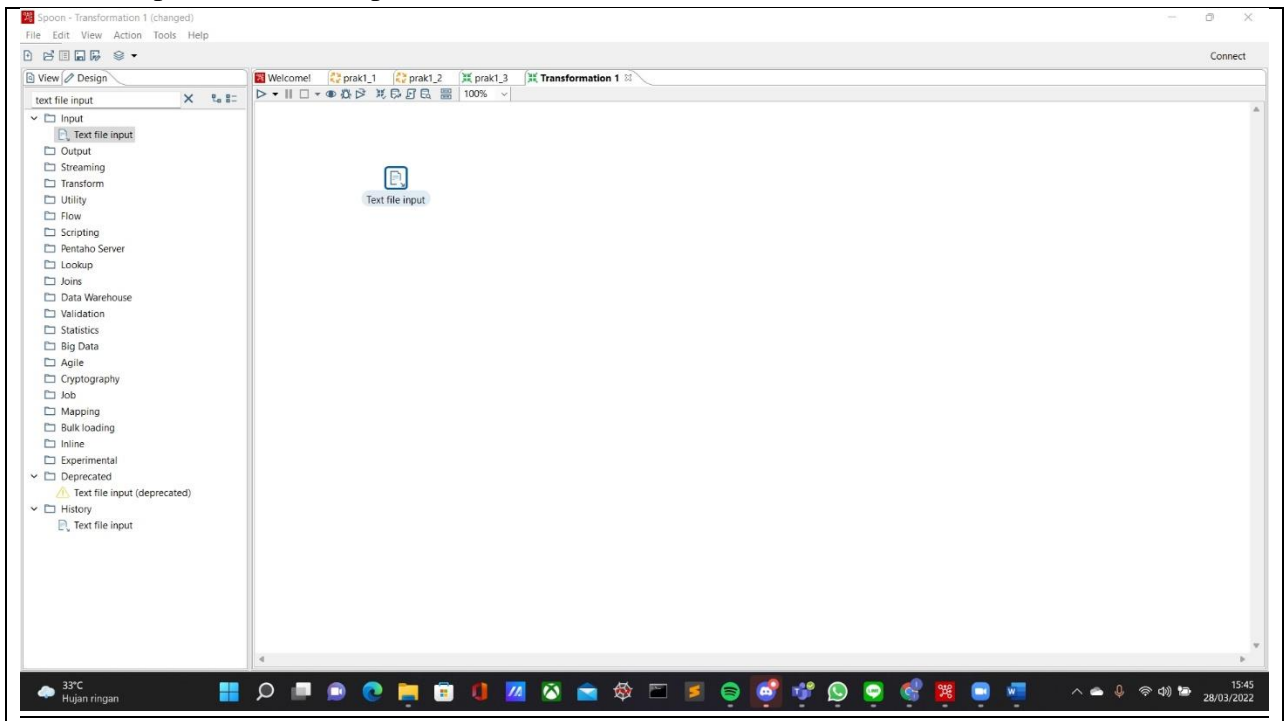
3. Elemen Kompetensi

a. Latihan pertama – Input Data Text (TXT)

1. Buat transformation sheer baru



2. Cari dan pilih Text File Input.



3. Klik 2x pada Text File Input lalu browse file yang ingin diinputkan, jika sudah setelah itu klik add. Maka tampilan sesudah di add akan seperti dibawah ini.

Text file input

Step name: Text file input

File | Content | Error Handling | Filters | Fields | Additional output fields

File or directory: D:\file kuliah smstr4\filepraktikdatawarehouse\Propinsi Di Pulau Jawa.txt [Add] [Browse...]

Regular Expression: []

Exclude Regular Expression: []

Selected files:

#	File/Directory	Wildcard (RegExp)	Exclude wildcard	Required	Include subfolders
1					

[Delete] [Edit]

Accept filenames from previous steps

Accept filenames from previous step ☐

Pass through fields from previous step ☐

Step to read filenames from: []

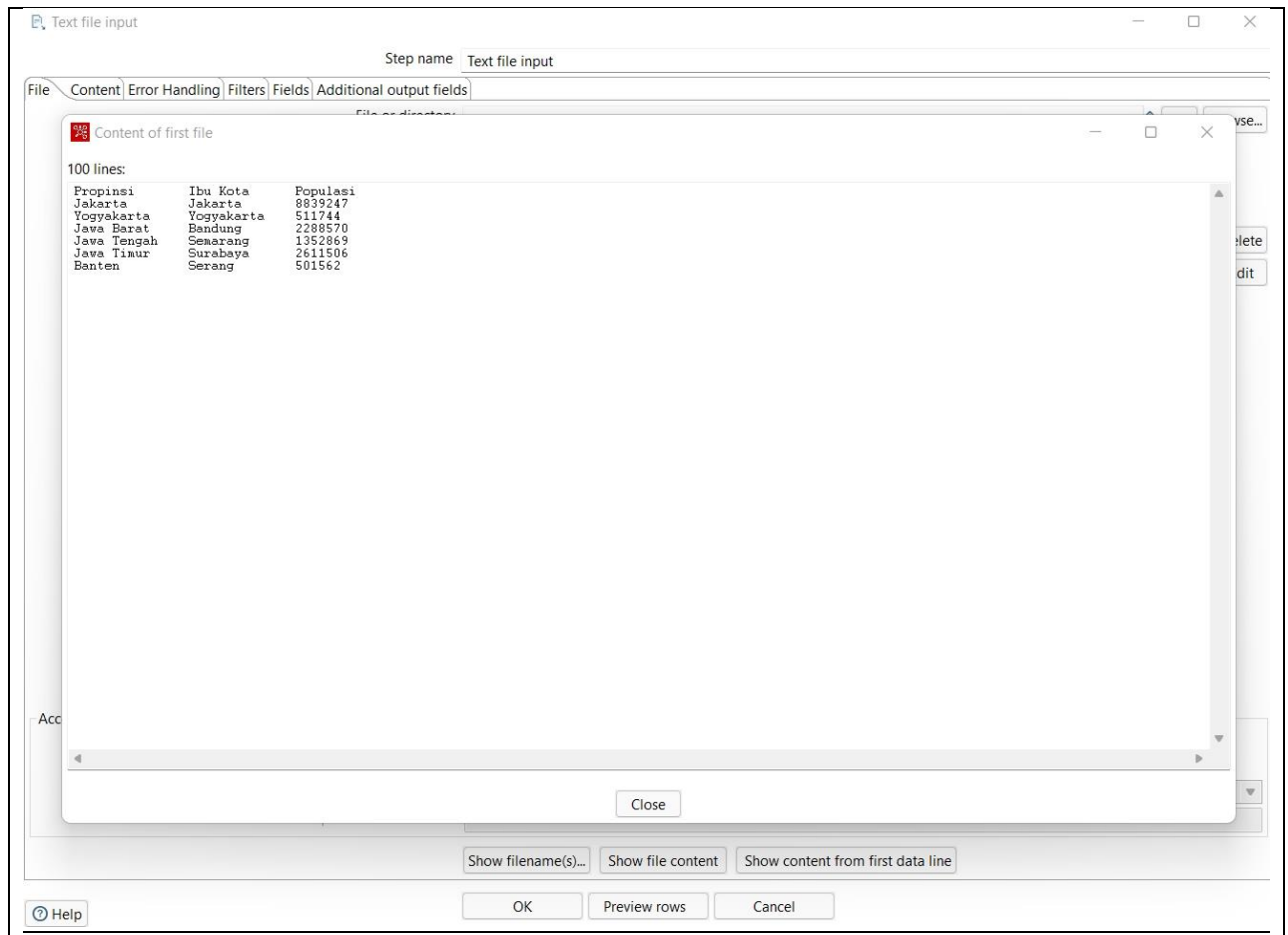
Field in the input to use as filename: []

[Show filename(s)...] [Show file content] [Show content from first data line]

[Help] [OK] [Preview rows] [Cancel]

4. Jika ingin melihat preview data, maka pilih show file content, lalu akan ditampilkan preview datanya seperti gambar dibawah ini





5. Jika sudah melakukan pemeriksaan data yang diinputkan setelah itu klik OK.



Text file input

Step name: Text file input

File | Content | Error Handling | Filters | Fields | Additional output fields

File or directory: Add Browse...

Regular Expression:

Exclude Regular Expression:

Selected files:

#	File/Directory	Wildcard (RegExp)	Exclude wildcard
1	D:\file kuliah smstr4\filepraktikdatawarehouse\Propinsi Di Pulau Jawa.txt		

Delete Edit

Accept filenames from previous steps

Accept filenames from previous step ☐

Pass through fields from previous step ☐

Step to read filenames from:

Field in the input to use as filename:

Show filename(s)... Show file content Show content from first data line

OK Preview rows Cancel

Help

Spoon - prak2_1

File Edit View Action Tools Help

View Design

Search

Input Output Streaming Transform Utility Flow Scripting Pentaho Server Lookup Joins Data Warehouse Validation Statistics Big Data Agile Cryptography Job Mapping Bulk loading Inline Experimental Deprecated History

Connect

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

2022/03/28 16:49:39 - Spoon - Running transformation using the Kettle execution engine

2022/03/28 16:49:39 - Spoon - Transformation opened.

2022/03/28 16:49:39 - Spoon - Launching transformation (prak2_1)...

2022/03/28 16:49:39 - Spoon - Started the transformation execution.

2022/03/28 16:49:39 - prak2_1 - Dispatching started for transformation (prak2_1)

2022/03/28 16:49:39 - Text file input.0 - Opening file: file:///D:/file%20kuliah%20smstr4/filepraktikdatawarehouse/Propinsi%20Di%20Pulau%20Jawa.txt

2022/03/28 16:49:39 - Text file input.0 - Finished processing (I=7, O=0, R=0, W=6, U=1, E=0)

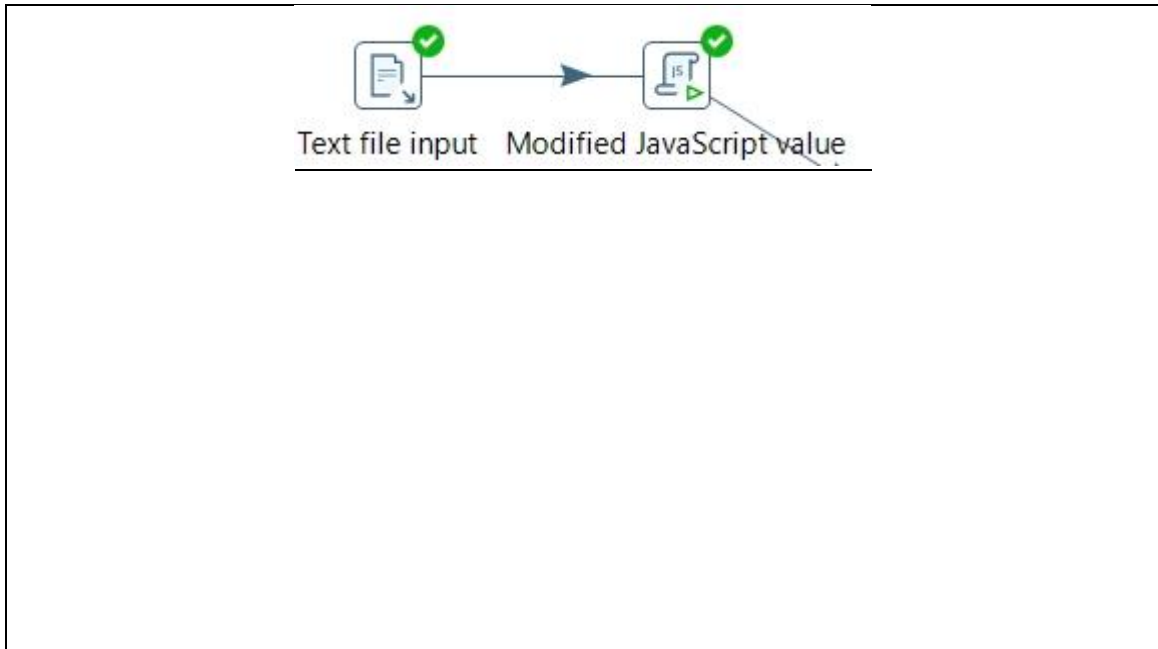
2022/03/28 16:49:39 - Spoon - The transformation has finished!!

32°C Hujan ringan 16:54 28/03/2022

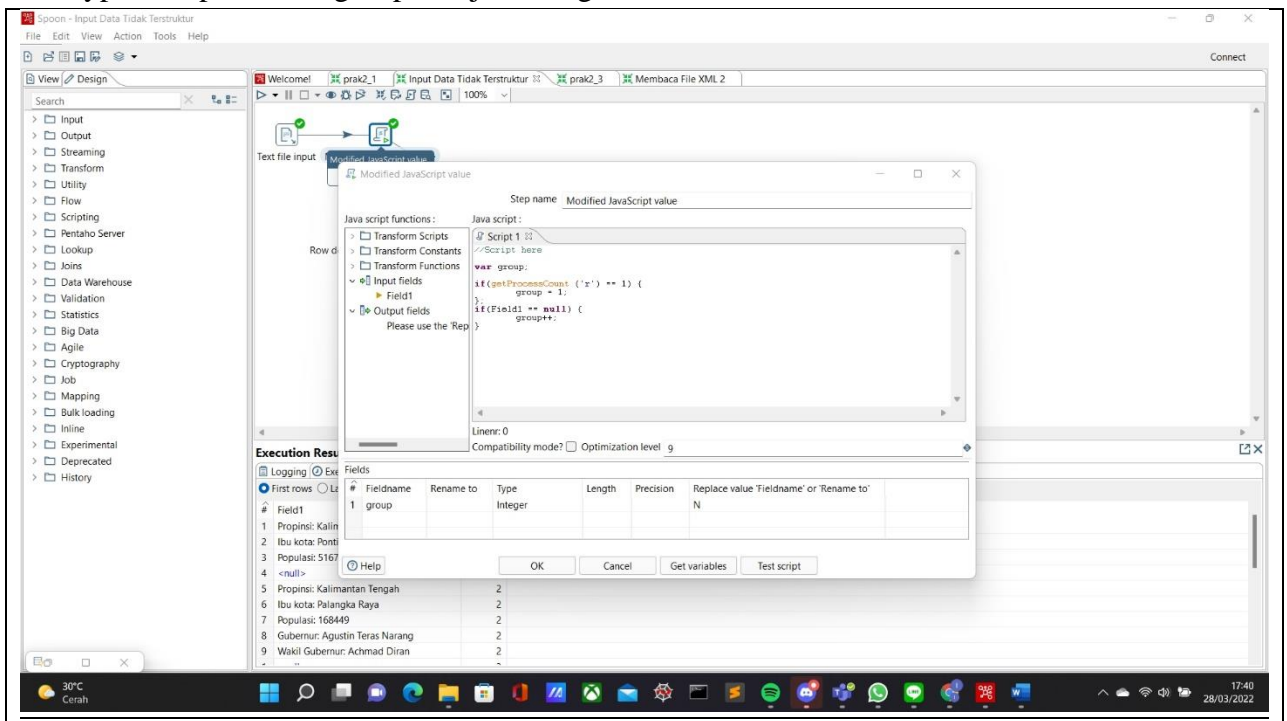


b. Latihan Kedua – Menginput Data Text yang Tidak Terstruktur

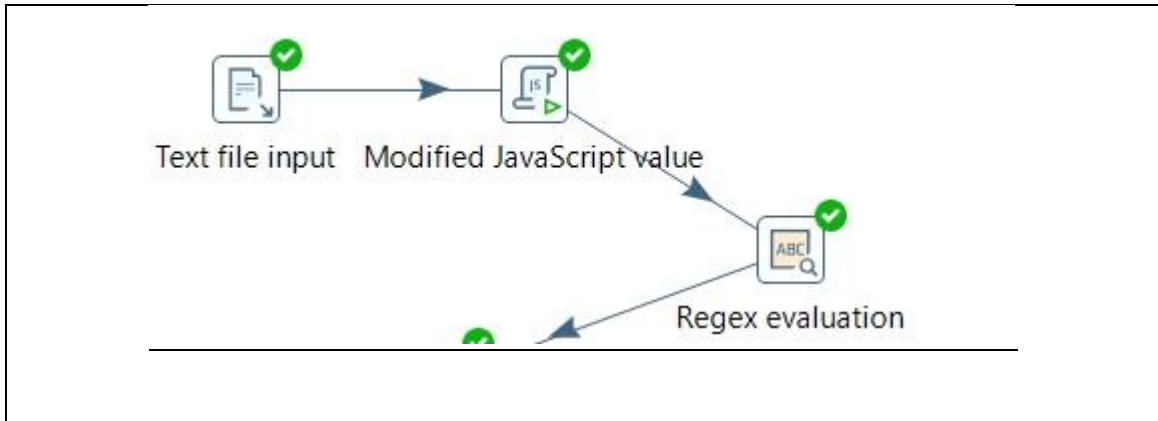
1. Ulangi semua langkah di **Latihan Pertama** dengan menginputkan data yang berbeda. dan setelah itu cari dan tambahkan Modified JavaScript value lalu hubungkan dengan Hop.



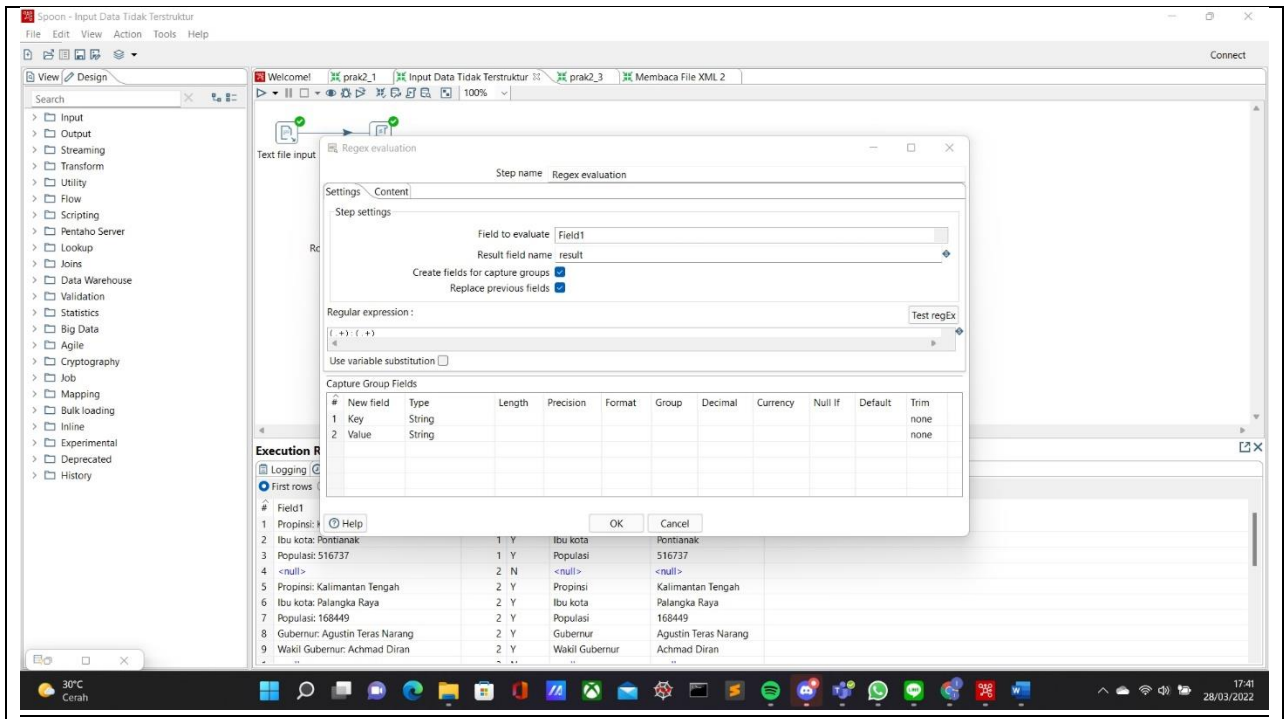
2. Selanjutnya tulis script seperti pada gambar dibawah dan klik Get variables setelah itu ubah type data pada field group menjadi Integer. Setelah itu klik OK.



3. Selanjutnya cari dan tambahkan Regex evaluation. Setelah itu klik 2x pada Regex Evaluation

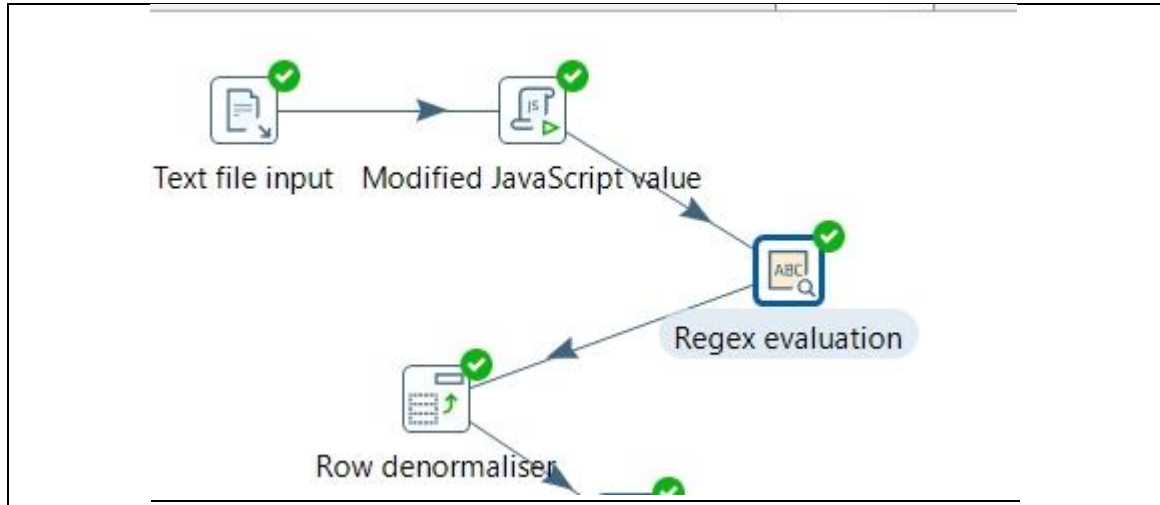


4. Konfigurasi bagian Regex Evaluation seperti gambar dibawah ini lalu setelah itu klik OK.

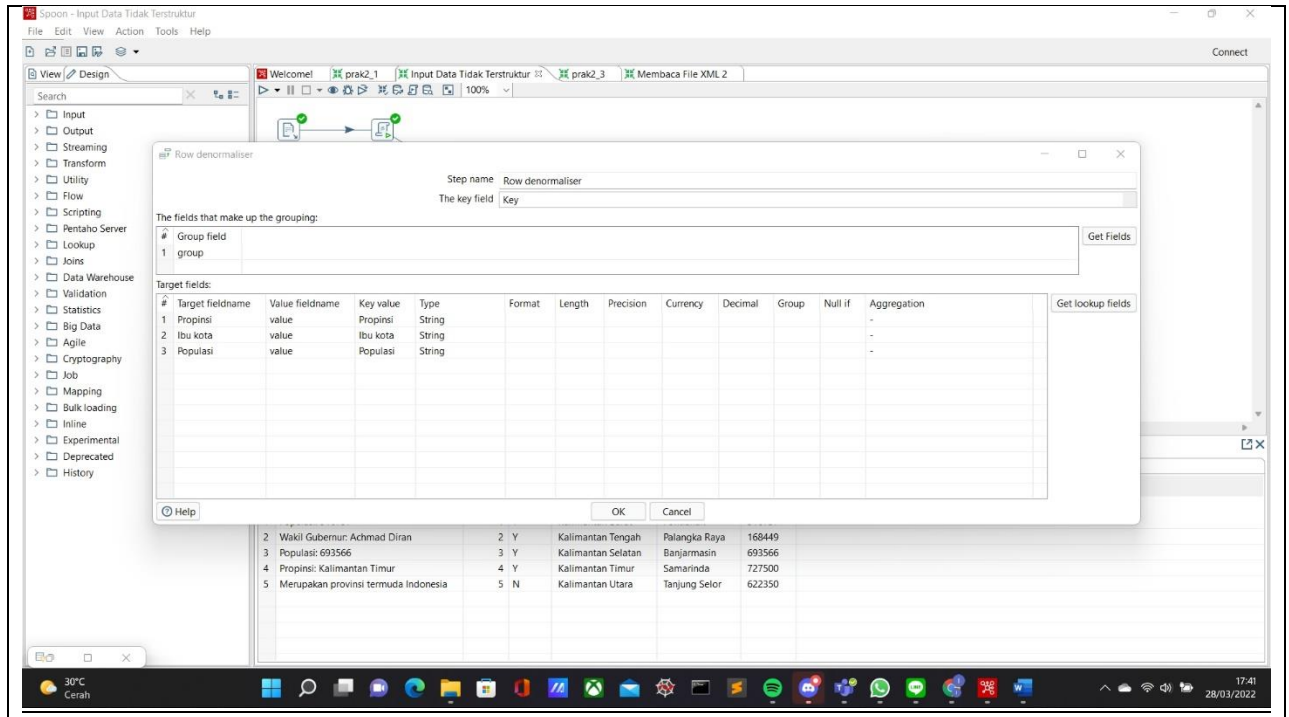


5. Setelah itu cari dan tambahkan Row denormaliser. Lalu klik 2x pada Row denormaliser yang sudah kita tambahkan.



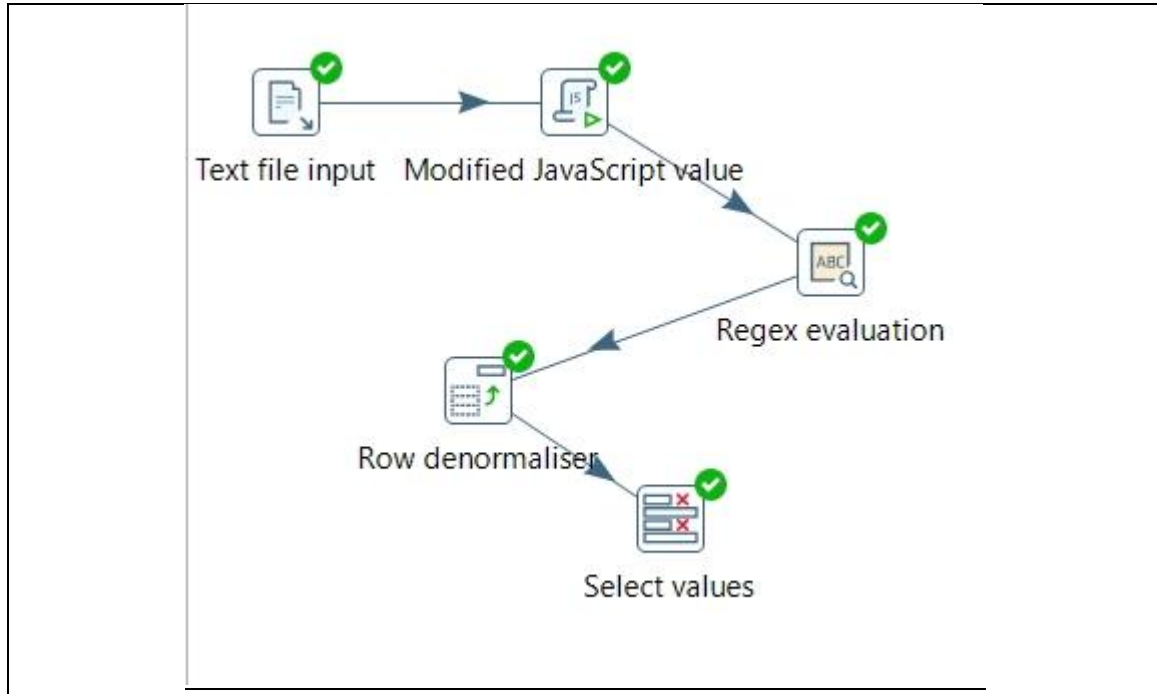


6. Konfigurasi Row normaliser seperti pada gambar dibawah ini lalu klik OK.



7. Setelah itu cari dan pilih Select values. Dan klik 2x pada Select values yang sudah ditambahkan.





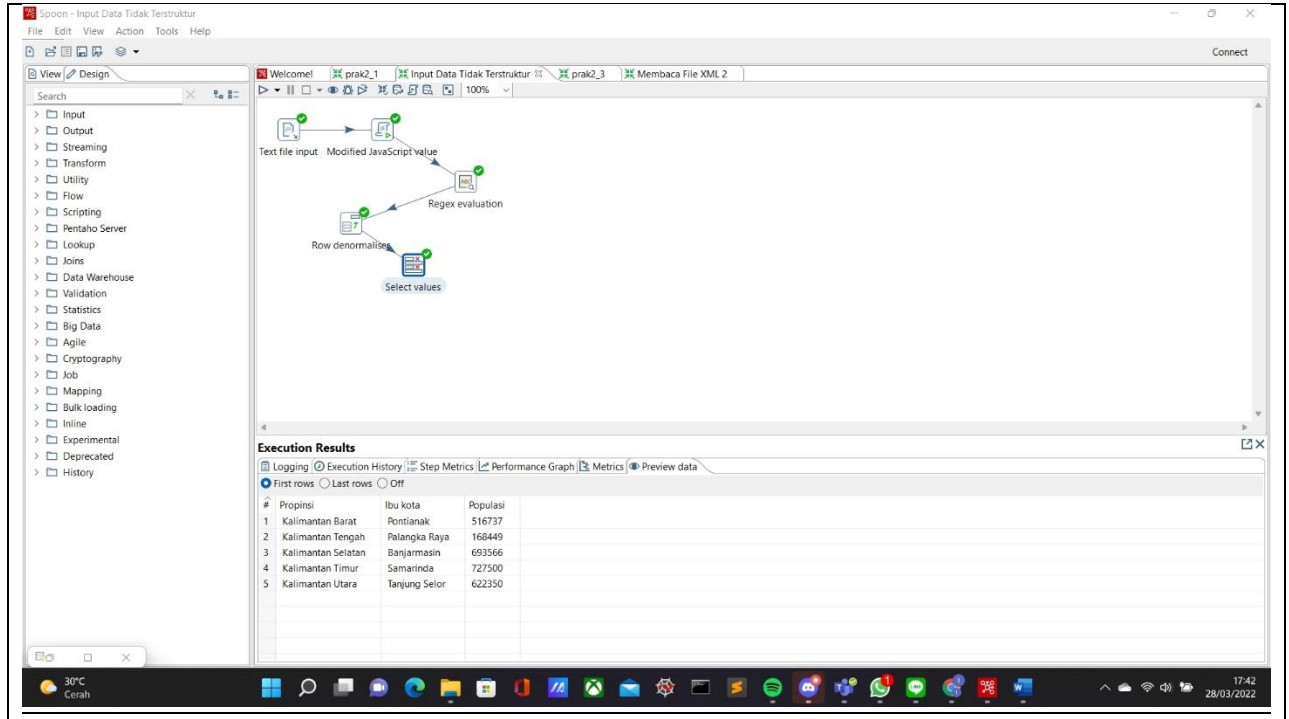
8. Setelah itu klik get values and hapus variabel yang tidak diperlukan dan sisakan variabel dibawah ini setelah itu klik OK. Setelah itu Jalankan/running

Execution Results

#	Propinsi	Ibu kota	Populasi
1	Kalimantan Barat	Pontianak	516737
2	Kalimantan Tengah	Palangka Raya	168449
3	Kalimantan Selatan	Bandarjaya	693566
4	Kalimantan Timur	Samarinda	727500
5	Kalimantan Utara	Tanjung Selor	622350

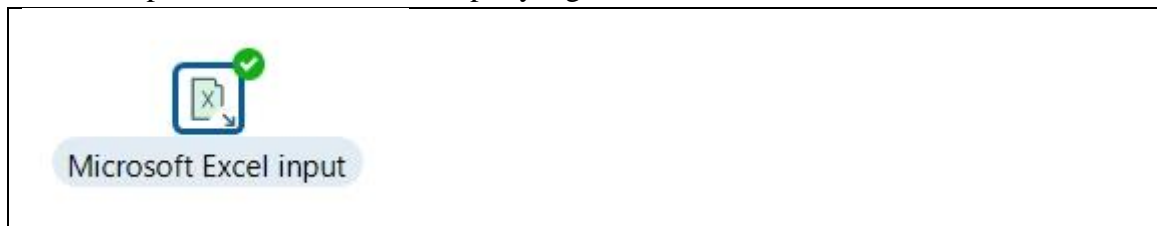
9. Berikut hasil setelah sudah dijalankan dan di running.





c. Latihan Ketiga – Menginput File Microsoft Excel

1. Buat transformation sheet baru dan cari serta tambahkan Microsoft Excel Input. Setelah itu klik 2x pada Microsoft Excel Input yang sudah ditambahkan tadi.



2. Browse serta add data yang ingin ditambahkan. Selanjutnya buka tab !Sheet.



Microsoft Excel input

Step name: Microsoft Excel input

Files | Sheets | Content | Error Handling | Fields | Additional output fields

Spread sheet type (engine): Excel 2007 XLSX (Apache POI)

File or directory:

Regular Expression:

Exclude Regular Expression:

Password:

Selected files:

#	File/Directory	Wildcard (RegExp)
1	D:\file kuliah smstr4\filepraktadawarehouse\Propinsi Di Sulawesi.xlsx	

Accept filenames from previous steps

Accept filenames from previous step ☐

Step to read filenames from:

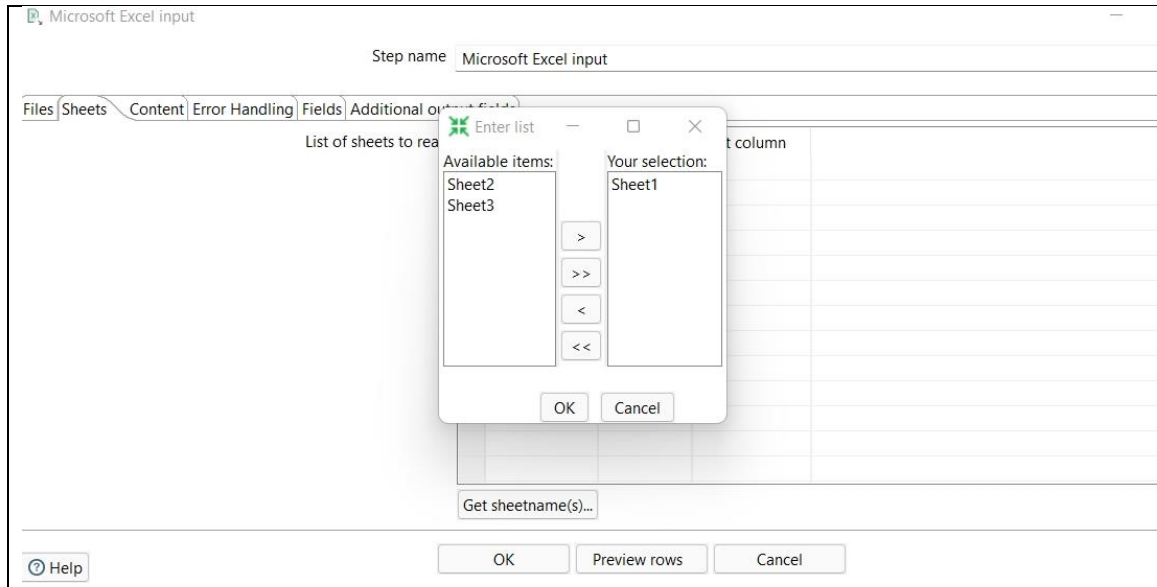
Field in the input to use as filename:

Show filename(s)...

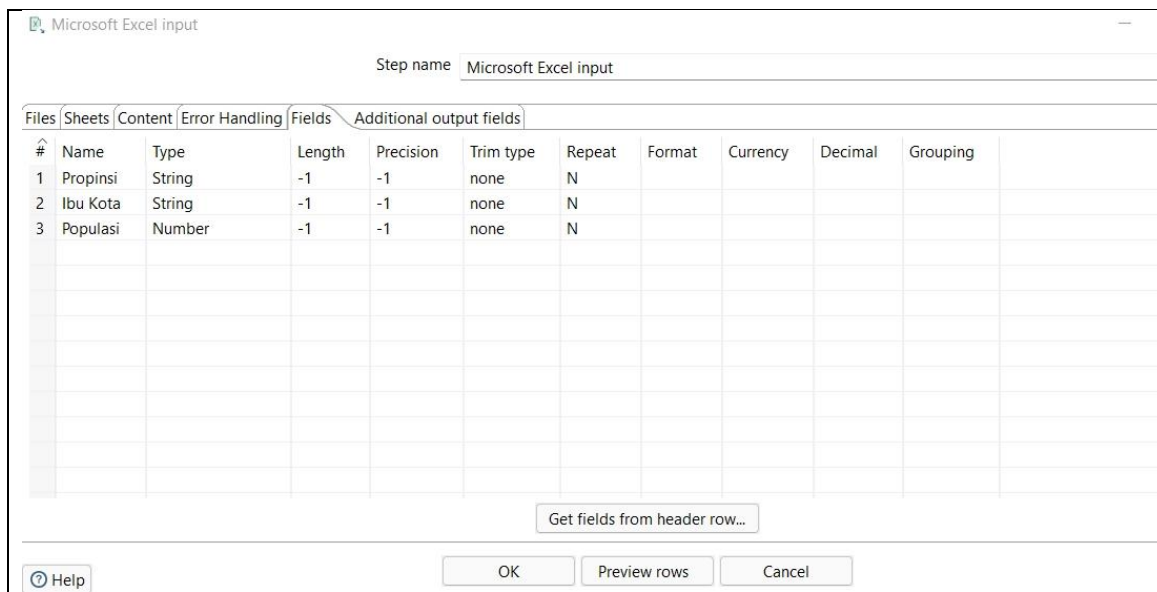
Help OK Preview rows Cancel

3. Klik get sheetsname lalu tambahkan Sheet yang berisi data (Sheet1) dan jika sudah klik OK dan pindah ke tab fields.

[illegible]

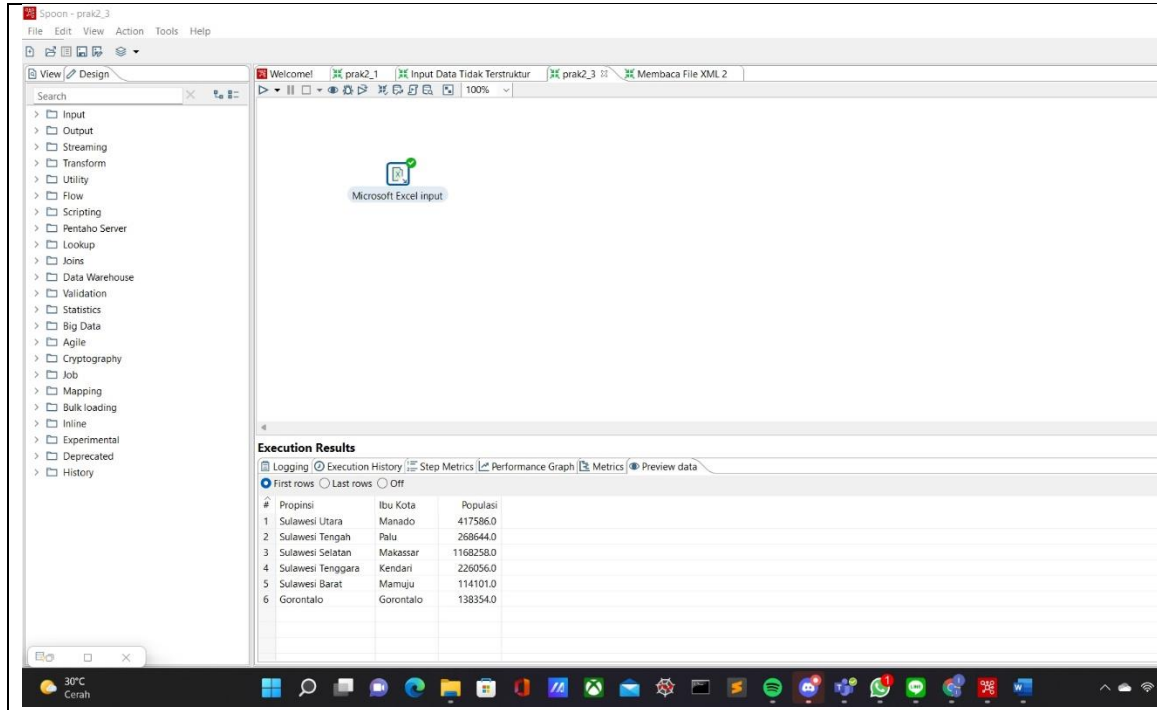


4. Lalu klik get fields from header row. Seperti pada gambar dibawah ini. Lalu setelah itu klik Preview rows.



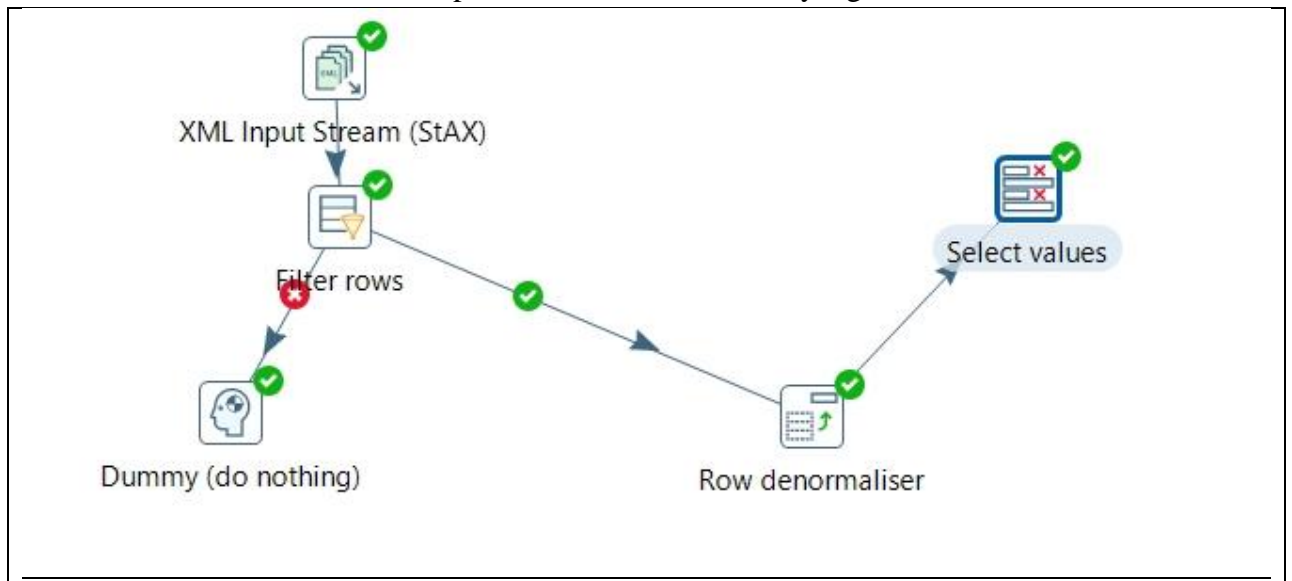
5. Berikut tampilan preview rows yang artinya data berhasil diinput. Jika sudah klik close dan klik OK.





d. Latihan Keempat – Menginput File XML

1. Buat transformation sheet baru dan cari serta tambahkan Get Data From XML pada transformation. Setelah itu klik 2x pada Get Data From XML yang sudah ditambahkan.



2. Browse dan cari data file XML yang ingin diinput lalu add atau tambahkan, setelah itu buka tab Fields.



XML input stream (StAX)

Step name XML Input Stream (StAX)

Filename \${Internal.Entry.Current.Directory}/Propinsi Lainnya.xml [Browse...](#)

Source is from a previous step ☐

Source field name

Add filename to result? ☐

Skip (Elements/Attributes) 0

Limit (Elements/Attributes) 0

Default String Length 1024

Encoding UTF-8

Add Namespace information? ☐

Trim strings? ☒

Include filename in output? ☐ Fieldname xml_filename

Row number in output? ☐ Fieldname xml_row_number

XML data type (numeric) in output? ☐ Fieldname xml_data_type_numeric

XML data type (description) in output? ☒ Fieldname xml_data_type_description

XML location line in output? ☐ Fieldname xml_location_line

XML location column in output? ☐ Fieldname xml_location_column

XML element ID in output? ☒ Fieldname xml_element_id

XML parent element ID in output? ☒ Fieldname xml_parent_element_id

XML element level in output? ☒ Fieldname xml_element_level

XML path in output? ☒ Fieldname xml_path

XML parent path in output? ☒ Fieldname xml_parent_path

XML data name in output? ☒ Fieldname xml_data_name

XML data value in output? ☒ Fieldname xml_data_value

[Help](#) [OK](#) [Preview](#) [Cancel](#)

3. Lalu klik Get Fields dan isi beberapa kolom data yang kosong seperti pada gambar dibawah. Lalu selanjutnya klik Preview rows.



The screenshot shows the SAP Data Services Designer interface. The design canvas displays a data flow starting with an 'XML Input Stream (StAX)' component, which connects to a 'Split rows' component. From 'Split rows', the flow goes to a 'Dummy (do nothing)' component and then to a 'Row normaliser' component. Finally, the 'Row normaliser' connects to a 'Select values' component. The 'Execution Results' panel at the bottom shows a table with 7 rows of data.

#	xml_data_type_description	xml_element_id	xml_parent_element_id	xml_element_level	xml_path	xml_parent_path	Propinsi	Ibu Kota	Populasi
1	CHARACTERS	5	2	3	/data/row/populasi	/data/row	Bali	Denpasar	491500
2	CHARACTERS	9	6	3	/data/row/populasi	/data/row	Nusa Tenggara Barat	Mataram	356141
3	CHARACTERS	13	10	3	/data/row/populasi	/data/row	Nusa Tenggara Timur	Kupang	269680
4	CHARACTERS	17	14	3	/data/row/populasi	/data/row	Maluku	Ambon	428585
5	CHARACTERS	21	18	3	/data/row/populasi	/data/row	Maluku Utara	Sofifi	163467
6	CHARACTERS	25	22	3	/data/row/populasi	/data/row	Papua	Jayapura	197396
7	CHARACTERS	29	26	3	/data/row/populasi	/data/row	Papua Barat	Manokwari	207398

4. Jika data yang diinputkan sudah berhasil ditampilkan maka klik close dan setelah itu klik OK.

This screenshot shows the same SAP Data Services Designer interface as the previous one, but the 'Execution Results' panel now displays a different set of data, likely after a refresh or re-execution. The table structure is the same, but the values for 'Propinsi', 'Ibu Kota', and 'Populasi' are different.

#	Propinsi	Ibu Kota	Populasi
1	Bali	Denpasar	491500
2	Nusa Tenggara Barat	Mataram	356141
3	Nusa Tenggara Timur	Kupang	269680
4	Maluku	Ambon	428585
5	Maluku Utara	Sofifi	163467
6	Papua	Jayapura	197396
7	Papua Barat	Manokwari	207398



4. File Praktikum

Github Repository:

<https://github.com/IbnuFajar7/Data-Warehouse/tree/main/Prak-2>

5. Soal Latihan

Soal:

1. Apa yang dimaksud Input Data?
2. Apa yang dimaksud data XML dan kenapa XML dijadikan sebagai format penyimpanan data?

Jawaban:

1. proses pemindahan data dari fisik menjadi digital yang dimana data tersebut akan diketik dan dimasukkan kedalam komputer.
2. XML adalah bahasa markup untuk keperluan umum yang disarankan oleh W3C untuk membuat dokumen markup keperluan pertukaran data antar sistem yang beraneka ragam. XML digunakan karna teks nya yang sederhana, format data akan tetap sama jika ditransfer ke server lain, mudah ditransfer ke OS baru

6. Kesimpulan

- a. Dalam pengerjaan praktikum Data Warehouse, kita harus benar-benar teliti dalam menginputkan suatu fungsi untuk menampilkan suatu keluaran pada layar dengan sesuai.
- b. Kita dapat mengetahui bahwa dalam praktikum kali ini ketelitian sangat diperhatikan, bagaimana penginputan model file, penginputan jenis file seperti apa yang harus digunakan, ketelitian nama dan beberapa source code akan sangat berpengaruh ke step-step selanjutnya.

7. Cek List (✓)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan Pertama	✓	
2.	Latihan Kedua	✓	
3.	Latihan Ketiga	✓	
4.	Latihan Keempat	✓	

8. Formulir Umpan Balik



No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan Pertama	10 Menit	1
2.	Latihan Kedua	10 Menit	1
3.	Latihan Ketiga	10 Menit	1
4.	Latihan Keempat	10 Menit	1

Keterangan:

1. Menarik
2. Baik
3. Cukup
4. Kurang

